Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 15 (2011) 3330 – 3334

Procedia Engineering

www.elsevier.com/locate/procedia

Advanced in Control Engineering and Information Science

An Improved Algorithm about Failure Agreement based on Distributed Shared Memory

Jun Wang [a], Yaqiong Jiang[b]

[a](Financial Department, Guilin University of Electronic Technology, China)*

[b](Network Information Center, Guilin University of Electronic Technology, China)

Abstract

According to the uniform addressing and direct localization of network address space, adopting multi-threaded, multi-copy store and block storage in pages, page tree sub-node parallel failure methods, these can improve the efficiency and parallel processing of the DSM system. It is significance for improving the performance of distributed systems.

Selection and/or peer-review under responsibility of [CEIS 2011]

Key words： Distributed shared memory； Failure Agreement； Algorithm； Security； Parallelism

1. Introduction

Distributed Shared Memory (Distributed Shared Memory: DSM) is used for the computer which do not share physical memory. It realizes an abstract of shared data. There are two methods for distributed data sharing： Replication and transfer [1]。 Process can read and update to access to DSM like accessing the common memory in address space . The purpose is to allow the system to use a shared memory programming model. One of the core problem is that how to get good performance when the system contains a large number of computers.

The access to DSM refer to the network communication [2]. However, it must have a system runned at the bottom to ensure its transparency. The transparency makes the data updating in various other

* Corresponding author: +8618977366928
E-mail address: wangjun@guet.edu.cn

doi:10.1016/j.proeng.2011.08.624

computers can be observed by process [3]。 The paging mechanism for application components within the process is transparent; they can read and write logically any data in the DSM.

It appears that the processes are accessing a single shared memory. In fact, the physical memory is distributed.With the development of multi-processing structure in shared-memory ,the capacity of computer main memory is expanding. Particularly in a distributed environment, the number of processors is up to hundreds or even thousands [4]， can effectively support parallel processing.The importance is increasing.It is an important research area.

2. Improvement ideas

The improve ideas are described from the protocol and algorithm in failure.

2.1. Failure

The algorithm uses page protection based on failure to achieve the data sharing. When there is a process updating a page, the process holds the data read and write permission on the local data. And other process do not have the permission to access to the page. When one or more processes to read a page, they only have the read permission. All other processes do not have access permission. In addition, there is no other case. The latest version of the page P has been designated as the process of the owner of this page - Owner (p). It may be a single write process, and it may be one of the reading processes. The process set which has a copy of the page P is called copyset (P).

There are three situation in system transition, and they are page failure, read failure and write failure. Page failure is that when a process P try to write to page, and it does not have the access permission or only have the read permission, the page failure will be happened in the system. When process P try to read page P, and it does not have the access permission to this page, the page failure will be happened in the system. Two or more processes which possess the read-only copy may have write failure at the same time.

2.2. Failure protocol and algorithm

Two important issues must be resolved in failure agreement. One is how to locate a given page p owner owner (p), and where to stored copyset (p).

The other one is how to effectively achieve the treatment of page failure.The following algorithm is an improved algorithm on central manager in the Ivy (Integrated shared Virtual memory at Yale) by Li and Hudak. The algorithm leave off a message on the basis of the original four messages[5]. In this algorithm, the system uses a single server which is called the manager to store the location of owner (p) (transfer address). The manager may be one of the process running application, and it could be other process. In this algorithm, the system will store copyset (p) to the owner (p) .It means that system store copyset (p) process of identification and transfer of members addresses.

When page failure happens, local process (called the customer client) should send messages containing page number and access type (read or read / write) to manager,and then the client waits for response. The manager processes the message through searching the address of owner (p) and sending the request to the page owner.Under the condition of write failure, manager will give the new ownership to the customer who writes data. The subsequent requests are queued on the client side until the customer completing the ownership conversed.

The previous owner passes this page to the customer. When the the event of write failure happened, it sends a copy set of the page copyset (p). When customers receive the copy of the set, it performs the failure process. It broadcasts a request to the member process of copyset, and waits for the confirmation

message of ‘failure has occurred’ sent by these processes. This broadcast does not need to sort the message, the client does not need to send the failure message to the former owner of pages.Because it is ineffective when the ownership is transfering.

In this improved algorithm about the central manager, the manager is a bottleneck in system performance, and its failure will cause the system to crash.

3. The improved algorithm

3.1. Improvement ideas

With the rapid development of computers, the Internet's bandwidth, speed, capacity is increasing. With the application of dual-core processor, parallelism and scalability are increased rapidly. With the growing of computer crime, security had been put on the agenda. With the development of electronic commerce and expanding of the scope of distributed systems, heterogeneity, openness, security, scalability, fault handling, concurrency and transparency are became the request for distributed system designers.It is much more complex than original algorithm that make more computers share the page load. But, based on the fixed distributed pages,using multicast and dynamic distributed management reflect flexibility, security, speed.

3.2. Algorithm and its description

The function of each site are the same as the function of manager in the improved central manager algorithm. Every site undertakes the task of fixed pages. The sites with the some page number are statically allocated a specific page. For example,the required data which is responsible for the function is stored in the site. The customer calculates the required number of pages, and then find the corresponding site in the address, which can improve the original algorithms focus on the management load. Of course, when the process is not balanced access to the page, the load of some manager is much bigger than that of other management. Compared with the original algorithm, the failure of a particular site do not cause the system to crash.

Based on this, the algorithm can solve the problem of network addressing using IPV6 .It adopt optimized method that is compatible with IPV4, such as IPV4, the multicast group specified by the D class of Internet addresses, the first 4 bits are the address of 1110, fast and flexible and have enough address space.

The location of page number is determined by the sequence of multicast. When a multicast message arrives at a computer, copy the message is delivered to local socket which is added to the specified multicast address and specified port number. Supposing that there are two clients C1, C2 to find a page possessed P by a process through multicasting, client C1 finds the P at first and obtains the ownership from P. It will record the sequence of multicast when the ownership of P is transferred. At the same time,the time stamp is modified the time for the transfer of ownership. C1 has the ownership of the page.The time stamp is also modified. When C2 multicast request arrives, because of the P do not possesses the page,and the time vector of C2 is after the time vector of p, P will refuse to C2 ,and put the C1 of copy page tree to C2. Now C2 applies the ownership to C1 directly. Because C1 has received the signal sent by C2,C1 will delete the expired request of C2 in the waiting queue (Delay is not lost for refusing to).After processing the operation itself,the ownership will be transferred to C2.What C1 should process are the same as what have be processed by the former owner P. After C2 finishing the operation, the content of the page updated will be back to P. While data are changed each time, new value will be broadcasted to every process of management copy data and the copy will be conserved.The better

performance can be obtained by the copy data[2]. Through adding C2 in the page directory tree of P, C3 can apply for a temporary possession of this page to C2,when site C3 has a failure request to P or P has some breakdown. When the P and C2 are normal working, you can pass page content to C3 adopting block way,sub-thread way to speed up the page content delivery speed. When the C3 just can read,a copy is added to P again.Even bandwidth consumption and processing time are increased, it also avoids data inconsistencies and loss phenomenon because of one page possessed by more processes.

For network security reasons, digital signatures and access control are increased. Digital signatures can effectively prevent false and unauthorized network access.That can increase the credibility, unforgetable, non-repudiation , processing targeted and eliminate unnecessary interference. The authority should be controlled. If the authority could be noen, read-only,read-write, three failure situation which are page failure, read failures and write failures in the original improved algorithm can be processed. In order to avoid page faults occurred frequently, broadcast page can be setted after 2n page failured. According to the system to set the size of the power n, if n equals 9 ,there are 512 Page failured,and the process of broadcasting to all pages of the current owner of the address in order to reduce addressing time.

In the original algorithm of original central manager, taking into the complexity of distributed systems, the pointer of process could point C2, C1, ... Cn from the original real owner. It leads the pointer chain growth.

In such an improved algorithm of dynamic and distributed management, when the ownership is transfered, directly two-way communication among sites can be executed in the condition of another process Cm requesting page P. That is because of the direct addressing and page directory tree are adopted. For example, real-time publisher / subscriber model and event-driven transmission can be adopted in order to achieve point to multi-point transmission. The pointer is shown as in Figure 5. It increases the flexibility of DSM system.

4. Conclusion

Through improving the algorithm about failure agreement, direct addressing to page can be achieved in process. That could make the process run directly in the document. It means that document space can be shared among each processes in the circumstance of distributed environment,that is also called distributed shared memory . Compared to the traditional operating system,it simplifies the structure of the operating system and it is convenient for user programming.The execution speed of process is improved.That is convenient for implementing process migration and realizing the load balancing of system.

Of course, as Kwlkarni etc pointed out: Different applications have different requirements. There is no single model can meet the needs of all applications.And there is no single system can be the best in all cases. In the design and implementation of the DSM, there is no single realization strategy or algorithm can be applied to all of the parallel computing problem. However, the proposed improved algorithm is undoubtedly of great significance to improve the performance of distributed systems.

References

[1] George Coulouris,Jean Dollimore,Tim Kindberg .Distributed Systems:Concepts and Design,Fourth Edition.[M]Beijing：Machinery Industry Press，2006.1

[2] Liu Fuyan,You Jinyuan.The achievement of distributed sharedmemory based on eliminating the process logical space [J]. Journal of Chinese computer systems,2001.03

[3] Andrew s.Tanenbaum.Distributed Operating Systems[M].Beijing：Machinery Industry Press，2006.07

[4] Peng Dechun，Qiu Yulan,et.al.The research on distributed shared memory [J]. Journal of Wuhan University,2006.10